OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 19:58:54 ; Search time 1988 Seconds
(without alignments)
244.512 Million cell updates/sec

Title: US-09-758-881-115
Perfect score: 20
Sequence: 1 gctccagcatctgctgcttc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 33330

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	13.4	67.0	29	28	AZ780164	AZ780164 2M0017E17
	2	12.2	61.0	25	9	AI748295	AI748295 sb50f02.y
c	3	12	60.0	24	28	AZ779573	AZ779573 2M0016K09
	4	12	60.0	27	28	AZ404206	AZ404206 1M0172120
c	5	12	60.0	29	29	TA141E08P	AL466622 T. brucei
c	6	11.8	59.0	24	29	BZ358821	BZ358821 SALK_1333
c	7	11.6	58.0	22	28	AZ976330	AZ976330 2M0251P08
	8	11.6	58.0	24	28	AZ826814	AZ826814 2M0102M15
	9	11.6	58.0	28	9	AI444428	AI444428 fb38c12.x
c	10	11.6	58.0	28	9	AI756191	AI756191 EtESTea40
c	11	11.2	56.0	28	9	AI755903	AI755903 EtESTea08
	12	11	55.0	22	28	AZ864977	AZ864977 2M0174D21
	13	11	55.0	28	9	AA717506	AA717506 vu22e03.r
	14	11	55.0	28	29	DME546620	AJ546620 Drosophil
	15	11	55.0	30	28	AZ396226	AZ396226 1M0160N09
c	16	11	55.0	30	28	BH810436	BH810436 SALK_0495
	17	11	55.0	30	28	BH847383	BH847383 SALK_0531
	18	10.8	54.0	19	28	AZ585898	AZ585898 1M0391L22
c	19	10.8	54.0	24	28	AZ394196	AZ394196 1M0157H08
c	20	10.8	54.0	25	28	AZ782142	AZ782142 2M0022H10
	21	10.8	54.0	29	28	AZ449708	AZ449708 1M0248M04
	22	10.6	53.0	19	28	AZ430028	AZ430028 1M0214L16
	23	10.6	53.0	21	28	AZ786813	AZ786813 2M0032K11
c	24	10.6	53.0	26	28	AZ657494	AZ657494 1M0533A15
	25	10.6	53.0	28	14	T47368	T47368 yb13b04.r1
	26	10.6	53.0	29	12	BM396481	BM396481 5009-0-21
	27	10.6	53.0	29	28	AZ854411	AZ854411 2M0158B05
	28	10.4	52.0	19	28	AZ613058	AZ613058 1M0441C18
c	29	10.4	52.0	20	28	AZ309949	AZ309949 1M0017M22
	30	10.4	52.0	22	29	PCH304213	AJ304213 Plasmodiu
	31	10.4	52.0	28	9	AI756191	AI756191 EtESTea40
	32	10.4	52.0	29	28	AZ769892	AZ769892 1M0571C03
c	33	10.2	51.0	19	28	AZ858978	AZ858978 2M0164F24
	34	10.2	51.0	25	9	AI594892	AI594892 ve12c06.x
c	35	10.2	51.0	25	9	AI743387	AI743387 wg92f01.x
	36	10.2	51.0	26	28	AZ620130	AZ620130 1M0452A14
c	37	10.2	51.0	28	12	BM399938	BM399938 5009-0-63
	38	10.2	51.0	30	10	BG424013	BG424013 602447475
c	39	10	50.0	22	28	AZ492512	AZ492512 1M0326K08
	40	10	50.0	24	28	AZ308017	AZ308017 1M0010M05
c	41	10	50.0	25	28	AZ364381	AZ364381 1M0110A08
c	42	10	50.0	26	28	AZ386258	AZ386258 1M0145E08
c	43	10	50.0	27	28	AZ797359	AZ797359 2M0053J06
	44	10	50.0	27	28	AZ876196	AZ876196 2M0191A12
	45	10	50.0	28	28	AZ596903	AZ596903 1M0410O24

ALIGNMENTS

RESULT 1
AZ780164/c
LOCUS AZ780164 29 bp DNA linear GSS 16-FEB-2001
DEFINITION 2M0017E17F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0017E17 F, genomic survey sequence.
ACCESSION AZ780164
VERSION AZ780164.1 GI:12911551
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia, Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 29)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel. 801 585 5606
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0017 row: E column: 17
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 29.

FEATURES Location/Qualifiers
source 1. .29
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0017E17"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 9 a 4 c 10 g 6 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 28; Length 29;
Best Local Similarity 93.3%; Pred. No. 1e+05;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CTCCAGCATCTGCTG 16
| ||||||||||||||
Db 27 CGCCAGCATCTGCTG 13

RESULT 2
AI748295
LOCUS AI748295 25 bp mRNA linear EST 30-NOV-2001
DEFINITION sb50f02.y1 Gm-c1011 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1011-340 5' similar to SW:GLC1_SOYBN P04776 GLYCININ G1 PRECURSOR [CONTAINS: GLYCININ A1A SUBUNIT; GLYCININ BX SUBUNIT]. ;, mRNA sequence.
ACCESSION AI748295
VERSION AI748295.1 GI:5126559
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids, eurosids I, Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 25)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com
Insert Length: 2068 Std Error: 0.00
High quality sequence stop: 1.

FEATURES Location/Qualifiers
source 1. .25
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1011-340"
/tissue_type="immature cotyledons of greenhouse grown plants"
/lab_host="DH10B"
/clone_lib="Gm-c1011"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2: XhoI; This cDNA library was constructed from mRNA isolated from immature cotyledons (100-200mgs) of greenhouse grown plants. The cDNA library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a Not I restrictions site. Sal I linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-Sal I restriction site of the pSPORT 1 vector. The ligated cDNA fragments were transformed into E. coli ElectroMax DH10B host cells. This library was constructed by Dr. Lila Vodkin and Dr. Anu Khanna."

BASE COUNT 2 a 5 c 4 g 14 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 25;
Best Local Similarity 82.4%; Pred. No. 2.8e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 TCCAGCATCTGCTGCTT 19
| ||| | |||||||||
Db 2 TTCAGTAGCTGCTGCTT 18

RESULT 3
AZ779573/c
LOCUS AZ779573 24 bp DNA linear GSS 16-FEB-2001
DEFINITION 2M0016K09F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0016K09 F, genomic survey sequence.
ACCESSION AZ779573
VERSION AZ779573.1 GI:12910362
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 24)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,P., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

```
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0016  row: K  column: 09
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 24.
FEATURES              Location/Qualifiers
     source           1. .24
                      /organism="Mus musculus"
                      /mol_type="genomic DNA"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC2M0016K09"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
BASE COUNT         8 a      8 c      8 g      0 t
ORIGIN

  Query Match             60.0%;  Score 12;  DB 28;  Length 24;
  Best Local Similarity   75.0%;  Pred. No. 3.3e+05;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GCTCCAGCATCTGCTGCTTC 20
              ||| | ||  |||||||| |
Db         23 GCTGCTGCTGCTGCTGCTGC 4


RESULT 4
AZ404206
LOCUS       AZ404206                 27 bp    DNA     linear   GSS 03-OCT-2000
DEFINITION  1M0172I20F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0172I20 F, genomic survey sequence.
ACCESSION   AZ404206
VERSION     AZ404206.1  GI:10528219
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg , 20 S  2030 E , SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0172  row: I  column: 20
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 27.
FEATURES              Location/Qualifiers
     source           1. .27
                      /organism="Mus musculus"
                      /mol_type="genomic DNA"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC1M0172I20"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
BASE COUNT         0 a      9 c      9 g      9 t
ORIGIN

  Query Match             60.0%;  Score 12;  DB 28;  Length 27;
  Best Local Similarity   75.0%;  Pred. No. 3.4e+05;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GCTCCAGCATCTGCTGCTTC 20
              ||| | ||  |||||||| |
Db          3 GCTGCTGCTGCTGCTGCTGC 22


RESULT 5
TA141E08P/c
LOCUS       TA141E08P                29 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 141e08, forward sequence,
            genomic survey sequence.
ACCESSION   AL466622
VERSION     AL466622.1  GI:11835977
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 29)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
```

```
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb). The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects. In
            Genome Sequencing: A Practical Approach, eds M Vaudin and B
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES                Location/Qualifiers
     source             1. .29
                        /organism="Trypanosoma brucei"
                        /mol_type="genomic DNA"
                        /strain="TREU927"
                        /db_xref="taxon.5691"
                        /clone-"141e08"
BASE COUNT         8 a      7 c      9 g      5 t
ORIGIN

  Query Match             60.0%;  Score 12;  DB 29;  Length 29;
  Best Local Similarity   75.0%;  Pred. No. 3.5e+05;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GCTCCAGCATCTGCTGCTTC 20
              | | | ||||| |||||| |
Db         22 GTTTCTGCATCCGCTGCTGC 3


RESULT 6
BZ358821/c
LOCUS       BZ358821                  24 bp    DNA     linear   GSS 14-NOV-2002
DEFINITION  SALK_133355.31.25.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_133355.31.25.x, genomic
            survey sequence.
ACCESSION   BZ358821
VERSION     BZ358821.1  GI:24951287
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A , Karnes,M , Kim,C.J , Parker,H , Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within an annotated exon of At4g02660.
            Class: TDNA tagged.
FEATURES                Location/Qualifiers
     source             1. .24
                        /organism="Arabidopsis thaliana"
                        /mol_type="genomic DNA"
                        /strain="Columbia 0"
                        /db_xref="taxon:3702"
                        /clone-"SALK_133355.31.25.x"
                        /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                        /note="PCR was performed on Arabidopsis thaliana lines
                        each of which contains one or more TDNA insertion
                        elements. The resultant fragment for each line was
                        directly sequenced to determine the genomic sequence at
                        the site of insertion. Details of the protocols used can
                        be found at http://signal.salk.edu/tdna_protocols.html"
BASE COUNT         9 a      5 c      6 g      4 t
ORIGIN

  Query Match             59.0%;  Score 11.8;  DB 29;  Length 24;
  Best Local Similarity   86.7%;  Pred. No. 3.9e+05;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 AGCATCTGCTGCTTC 20
              ||| | |||||||||
Db         24 AGCTTGTGCTGCTTC 10


RESULT 7
AZ976330/c
LOCUS       AZ976330                  22 bp    DNA     linear   GSS 27-APR-2001
DEFINITION  2M0251P08R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
            clone UUGC2M0251P08 R, genomic survey sequence.
ACCESSION   AZ976330
VERSION     AZ976330.1  GI:13847557
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm 308, Biomedical Polymers Research Bldg , 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000  Std Error: 0.00
            Plate: 0251  row: P  column: 08
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 22.
FEATURES                Location/Qualifiers
     source             1. .22
                        /organism="Mus musculus"
                        /mol_type="genomic DNA"
                        /strain="C57BL/6J"
                        /db_xref="taxon:10090"
                        /clone="UUGC2M0251P08"
                        /sex="Female"
                        /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
                        /clone_lib="Mouse 10kb plasmid UUGC2M library"
                        /note="Vector: PWD42nv, Purified genomic DNA from M.
                        musculus C57BL/6J (female) was obtained from the Jackson
                        Laboratory Mouse DNA Resource
                        (http://www.jax.org/resources/documents/dnares/). The DNA
                        was hydrodynamically sheared by repeated passage through a
                        0.005 inch orifice at constant velocity. The sheared DNA
                        was blunt end-repaired with T4 DNA polymerase and T4
                        polynucleotide kinase. Adaptor oligonucleotides were
                        ligated to the blunt ends in high molar excess. The
                        adaptored DNA was purified and size-selected for a 9.5 to
                        10.5 kb range using preparative agarose gel
                        electrophoresis. Vector DNA was prepared from a derivative
                        of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                        inducible derivative of plasmid R1. The vector was ligated
                        with adaptors complementary to the insert adaptors and
                        purified. The sheared, adaptored mouse DNA was annealed to
```

adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT 11 a 2 c 6 g 3 t
ORIGIN

Query Match 58.0%; Score 11.6; DB 28; Length 22;
Best Local Similarity 77.8%; Pred No. 4.6e+05;
Matches 14; Conservative 0; Mismatches 4, Indels 0, Gaps 0,

Qy 2 CTCCAGCATCTGCTGCTT 19
|| |||||||| | |||
Db 19 CTTCAGCATCTTTTCCTT 2

RESULT 8
AZ826814
LOCUS AZ826814 24 bp DNA linear GSS 20-FEB-2001
DEFINITION 2M0102M15R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0102M15 R, genomic survey sequence.
ACCESSION AZ826814
VERSION AZ826814.1 GI:12996722
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 24)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0102 row: M column: 15
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 24.
FEATURES Location/Qualifiers
source 1. .24
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0102M15"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and

purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT 1 a 9 c 5 g 9 t
ORIGIN

Query Match 58.0%; Score 11.6; DB 28; Length 24;
Best Local Similarity 77.8%; Pred. No. 4.7e+05;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 TCCAGCATCTGCTGCTTC 20
|||||| ||| |||| |
Db 1 TCCAGCCCCTGTTGCTGC 18

RESULT 9
AI444428
LOCUS AI444428 28 bp mRNA linear EST 07-JUN-2001
DEFINITION fb38c12.x1 Zebrafish WashU MPIMG EST Danio rerio cDNA clone IMAGE:3714166 3' similar to SW:RL5_RAT P09895 60S RIBOSOMAL PROTEIN L5. ;, mRNA sequence.
ACCESSION AI444428
VERSION AI444428.1 GI:4281620
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 28)
AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,P., Waterston,R. and Wilson,R.
TITLE WashU Zebrafish EST Project 1998
JOURNAL Unpublished
COMMENT Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
cDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by: Matthew Clark. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Genome Systems, St. Louis, Missouri (web address: www.genomesystems.com) (email contact: info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address: www.rzpd.de)
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: T7 ET from Amersham
High quality sequence stop: 1
POLYA=No.
FEATURES Location/Qualifiers
source 1. .28
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:3714166"
/sex="mixed"
/tissue_type="26 somite embryos, adult livers, shield stage embryos"
/lab_host="XL1-blue MRF"
/clone_lib="Zebrafish WashU MPIMG EST"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st strand cDNA was primed with a Not I - oligo(dT)15 primer [5'pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT3']; double-stranded cDNA was ligated to Sal I adaptors (BRL),

digested with Not I and cloned into the Not I and Sal I sites of the pSPORT1 vector (BRL). Library was constructed by Matthew Clark (Lehrach lab; ICRF, London and Max Planck Institut fuer Molekulare Genetik,Berlin). cDNAs for EST analysis were selected following oligonucleotide hybridization fingerprinting of arrayed clones from zebrafish late somitogenesis (26 ss), adult liver or embryonic shield stage (5.6 h) libraries. Fingerprint data were used to computationally cluster cDNAs, and a single cDNA from each cluster was chosen for sequencing. In some cases multiple members of the same cluster were sequenced to assess clustering parameters or single clones were sequenced additional times to assess quality control."

```
BASE COUNT        4 a      11 c       4 g       9 t
ORIGIN

  Query Match             58.0%;  Score 11.6;  DB 9;  Length 28;
  Best Local Similarity   77.8%;  Pred. No. 5e+05;
  Matches   14;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 CTCCAGCATCTGCTGCTT 19
              ||||| |  ||||| |||
Db         10 CTCCATCCACTGCTCCTT 27
```

RESULT 10
AI756191/c

LOCUS	AI756191 28 bp mRNA linear EST 18-JAN-2000
DEFINITION	EtESTea40h07 y1 Eimeria S5-2 Sporozoite stage Eimeria tenella cDNA 5' similar to TR:Q64507 Q64507 SERINE 1 ULTRA HIGH SULFUR PROTEIN. ;, mRNA sequence.
ACCESSION	AI756191
VERSION	AI756191.1 GI:5149914
KEYWORDS	EST.
SOURCE	Eimeria tenella
ORGANISM	Eimeria tenella Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae; Eimeria.
REFERENCE	1 (bases 1 to 28)
AUTHORS	Liberator,P., Diaz,C., Tang,K., Marra,M., Hillier,L., Kucaba,T., Martin,J., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen ,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey ,N., Schurk,R., Ritter,E., Kohn,S., Florence,N., Shin,T., Jackson ,Y., Cardenas,M., McCann,R., Waterston,R., Wilson,R. and Sibley,D.
TITLE	WashU-Merck Eimeria tenella project
JOURNAL	Unpublished
COMMENT	Contact: David Sibley, Ph.D. WashU-Merck Eimeria tenella project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu Contact David Sibley (toxoest@borcim.wustl.edu) for further information relating to organism, libraries, or clone availability. Trace considered overall poor quality Possible reversed clone: similarity on wrong strand Seq primer: -40RP from Gibco High quality sequence stop: 1.

FEATURES	Location/Qualifiers
source	1. .28 /organism="Eimeria tenella" /mol_type="mRNA" /strain="LS18" /db_xref="taxon:5802" /dev_stage="Sporozoite" /lab_host="SOLR E. coli" /clone_lib="Eimeria S5-2 Sporozoite stage" /note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI ; Sporozoites were obtained from in vitro sporulated and excysted oocysts of E. tenella grown in chickens. cDNA was synthesized from poly mRNA using an oligo-dT primer containing a XhoI site. Following second strand synthesis, EcoRI adapters were ligated to the cDNA and products were size-selected on Sephacryl S500. cDNAs were digested with EcoRI/XhoI and cloned into lambda Zap II (Stratagene). Clones were converted to phagemids by mass excision using ExAssist helper phage and SOLR cells (Stratagene). Insert sizes range from 1.2-2.9 kb."

```
BASE COUNT        6 a      11 c       9 g       2 t
ORIGIN

  Query Match             58.0%;  Score 11.6;  DB 9;  Length 28;
  Best Local Similarity   77.8%;  Pred. No. 5e+05;
  Matches   14;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GCTCCAGCATCTGCTGCT 18
              ||| | ||| | ||||||
Db         18 GCTGCCGCAGCAGCTGCT 1
```

RESULT 11
AI755903/c

LOCUS	AI755903 28 bp mRNA linear EST 18-JAN-2000
DEFINITION	EtESTea08h09.y1 Eimeria M5-6 Merozoite stage Eimeria tenella cDNA 5' similar to WP:T01B7.8 CE03592 ,, mRNA sequence.
ACCESSION	AI755903
VERSION	AI755903.1 GI:5149626
KEYWORDS	EST.
SOURCE	Eimeria tenella
ORGANISM	Eimeria tenella Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae; Eimeria.
REFERENCE	1 (bases 1 to 28)
AUTHORS	Liberator,P., Diaz,C., Tang,K., Marra,M., Hillier,L., Kucaba,T., Martin,J., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen ,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey ,N., Schurk,R., Ritter,E., Kohn,S., Florence,N., Shin,T., Jackson ,Y., Cardenas,M., McCann,R., Waterston,R., Wilson,R. and Sibley,D.
TITLE	WashU-Merck Eimeria tenella project
JOURNAL	Unpublished
COMMENT	Contact: David Sibley, Ph.D. WashU-Merck Eimeria tenella project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu Contact David Sibley (toxoest@borcim.wustl.edu) for further information relating to organism, libraries, or clone availability. Trace considered overall poor quality Possible reversed clone: similarity on wrong strand Seq primer: -40RP from Gibco High quality sequence stop: 1.

FEATURES	Location/Qualifiers
source	1. .28 /organism="Eimeria tenella" /mol_type="mRNA" /strain="LS18" /db_xref="taxon:5802" /dev_stage="Merozoite" /lab_host="SOLR E. coli" /clone_lib="Eimeria M5-6 Merozoite stage" /note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI ; Merozoites were obtained from ceacal scrapings of chickens infected with E. tenella. The library may contain a small percentage of host or bacterial contaminents. cDNA was synthesized from poly mRNA using an oligo-dT primer containing a XhoI site. Following second strand synthesis, EcoRI adapters were ligated to the cDNA and products were size-selected on Sephacryl S500. cDNAs were digested with EcoRI/XhoI and cloned into lambda Zap II (Stratagene). Clones were converted to phagemids by mass excision using ExAssist helper phage and SOLR cells (Stratagene). Insert sizes range from 0.7-1.5

kb."
BASE COUNT 8 a 7 c 9 g 4 t
ORIGIN

Query Match 56.0%; Score 11.2; DB 9; Length 28;
Best Local Similarity 81.2%; Pred. No. 7e+05;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GCTCCAGCATCTGCTG 16
|||| ||| ||||||
Db 21 GCTCAAGCTGCTGCTG 6

RESULT 12
AZ864977
LOCUS AZ864977 22 bp DNA linear GSS 21-FEB-2001
DEFINITION 2M0174D21R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0174D21 R, genomic survey sequence.
ACCESSION AZ864977
VERSION AZ864977.1 GI:13064817
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 22)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0174 row: D column: 21
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 22.
FEATURES Location/Qualifiers
source 1. .22
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0174D21"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 3 a 7 c 5 g 7 t
ORIGIN

Query Match 55.0%; Score 11; DB 28; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.7e+05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 CATCTGCTGCT 18
||||||||||
Db 2 CATCTGCTGCT 12

RESULT 13
AA717506
LOCUS AA717506 28 bp mRNA linear EST 29-DEC-1997
DEFINITION vu22e03.r1 Barstead mouse myotubes MPLRB5 Mus musculus cDNA clone IMAGE:1181404 5' similar to TR:Q60961 Q60961 GOLGI 4-TRANSMEMBRANE SPANNING TRANSPORTER MTP. ;, mRNA sequence.
ACCESSION AA717506
VERSION AA717506.1 GI:2729780
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 28)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:639252
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .28
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C3H"
/db_xref="taxon:10090"
/clone="IMAGE:1181404"
/cell_line="C2C12"
/lab_host="DH10B"
/clone_lib="Barstead mouse myotubes MPLRB5"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [AATTCGGATCCTTG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library constructed by Bob Barstead. The C2C12 cell line (available from ATCC, catalog # CRL-1772) differentiates rapidly, forming contractile myotubes and producing characteristic muscle proteins. "
BASE COUNT 4 a 11 c 8 g 5 t
ORIGIN

```
  Query Match             55.0%;  Score 11;  DB 9;  Length 28;
  Best Local Similarity   73.7%;  Pred. No. 8.4e+05;
  Matches   14;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 CTCCAGCATCTGCTGCTTC 20
              || ||||| | | |||| |
Db         10 CTACAGCACCCGGTGCTGC 28


RESULT 14
DME546620
LOCUS       DME546620                 28 bp    DNA     linear   GSS 24-FEB-2003
DEFINITION  Drosophila melanogaster flanking sequence of RS P element insertion
            P{RS5}5-SZ-3275, clone library P{RS5}, genomic survey sequence.
ACCESSION   AJ546620
VERSION     AJ546620.1  GI:28554721
KEYWORDS    GSS; genome survey sequence.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda, Hexapoda, Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1
  AUTHORS   Ryder,E.J., Ashburner,M., Bagunya,J., Blows,F., Bucheton,A.,
            Coulson,D., Dickson,B., Drummond,J., Glover,D., Gunton,N.,
            Hafen,E., Hall,S., Heisenberg,M., Lepesant,J.A., Maroy,P.,
            Mechler,B., O'Kane,C., Pflugfelder,G., Rasmuson-Lestander,A.,
            Reuter,G., Roote,J., Szidonya,J., Wang,S., Webster,J. and
            Russell,S.
  TITLE     Mapping of RS P element insertions in Drosophila melanogaster for
            the DrosDel second generation deficiency kit
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 28)
  AUTHORS   Ryder,E.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-FEB-2003) Ryder E.J., Department of Genetics,
            University of Cambridge, Downing Street, CB2 3EH, UNITED KINGDOM
COMMENT     The insertion point of the P element is before base 1 of the
            sequence. Further information about this P element insertion line
            can be found at http://www.flyseq.org.uk and
            http://www.drosdel.org.uk.
FEATURES             Location/Qualifiers
     source          1. .28
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /chromosome="3R"
                     /clone="P{RS5}5-SZ-3275"
                     /clone_lib="P{RS5}"
                     /note="read=5' end"
     misc_feature    1. .28
                     /note="P element insertion in the 3' to 5' orientation"
BASE COUNT        1 a     12 c      5 g     10 t
ORIGIN

  Query Match             55.0%;  Score 11;  DB 29;  Length 28;
  Best Local Similarity   73.7%;  Pred. No. 8.4e+05;
  Matches   14;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 CTCCAGCATCTGCTGCTTC 20
              |||   |  ||||||||||
Db          1 CTCTCCCGACTGCTGCTTC 19


RESULT 15
AZ396226
LOCUS       AZ396226                  30 bp    DNA     linear   GSS 03-OCT-2000
DEFINITION  1M0160N09R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0160N09 R, genomic survey sequence.
ACCESSION   AZ396226
VERSION     AZ396226.1  GI:10511298
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa, Chordata, Craniata, Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000    Std Error: 0.00
            Plate: 0160  row: N  column: 09
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 30.
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0160N09"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        6 a      7 c      5 g     12 t
ORIGIN

  Query Match             55.0%;  Score 11;  DB 28;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 8.6e+05;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 AGCATCTGCTG 16
              ||||||||||
Db          1 AGCATCTGCTG 11


Search completed: August 19, 2003, 21:20:41
Job time : 1993 secs
```